

Report 4: COVID-19: analysis of two serological tests

Davron Khamidov, s273331,

ICT for Health attended in A.Y. 2021/22

December 8th, 2021

1 Introduction

Spreading of COVID-19 (COrona VIRus Disease 19) infection can be reduced with early detection of infected people, so that they can start quarantine as soon as possible. The nasopharyngeal swab test is highly reliable but it requires time and is expensive, serological tests are faster and cheaper, but less reliable. Serological tests find the presence of IgG (Immunoglobulin G) and a high level of this antibody in blood means that the person is or has been affected by COVID-19.

This work reports the results of the analysis of two serological tests, discussing the setting of the thresholds to declare a positive result.

2 Method

A group of 879 people was subjected to 3 tests: one nasopharyngeal swab test and two serological tests (Test 1 and Test 2 in the following), recording the amount of IgG; 17 cases were removed from the dataset due to an uncertain swab test result. The positive swab tests were 71, whereas the negative ones were 791.

Test 1 contained 47 outliers, which were identified using DBSCAN [1] with parameters $\epsilon = 2.8$ and $M = 13$, and then removed (only from Test 1). In these cases, the swab test consisted of a combination of positive and negative results.

Swab test result was considered correct, and ROC curve (sensitivity versus false alarm, see Fig. 1) was measured for the two serological tests. The area under ROC was measured equal to 0.926 for Test 1 and 0.928 for Test 2.

For convenience, sensitivity and specificity versus threshold are also plotted in Fig. 2.

The following notation will be used: D means that the patient is really infected, H means that the patient is healthy, T_p means that the test is positive, T_n means that the test is negative.

The general approach in case of a positive serological test is to check again the person using the nasopharyngeal swab. This makes acceptable a relatively large false positive prob-

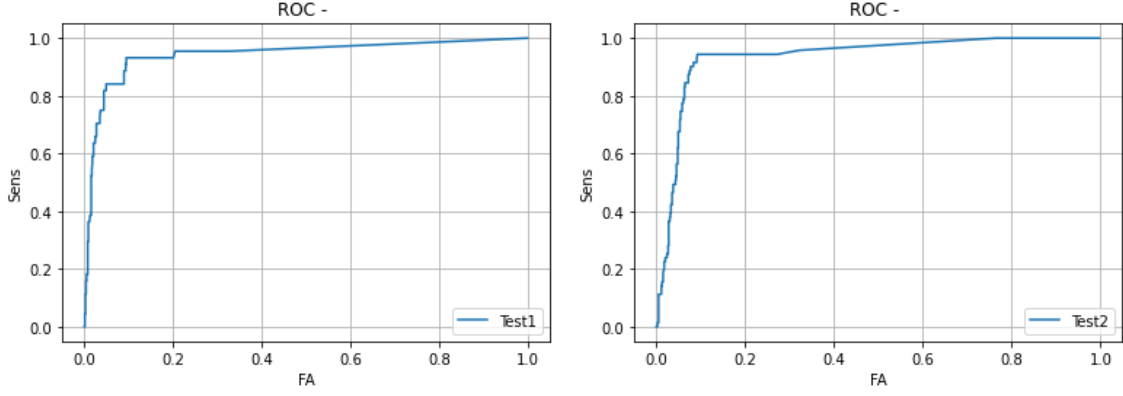


Figure 1: ROC curve for Test 1 (left) and Test 2 (right).

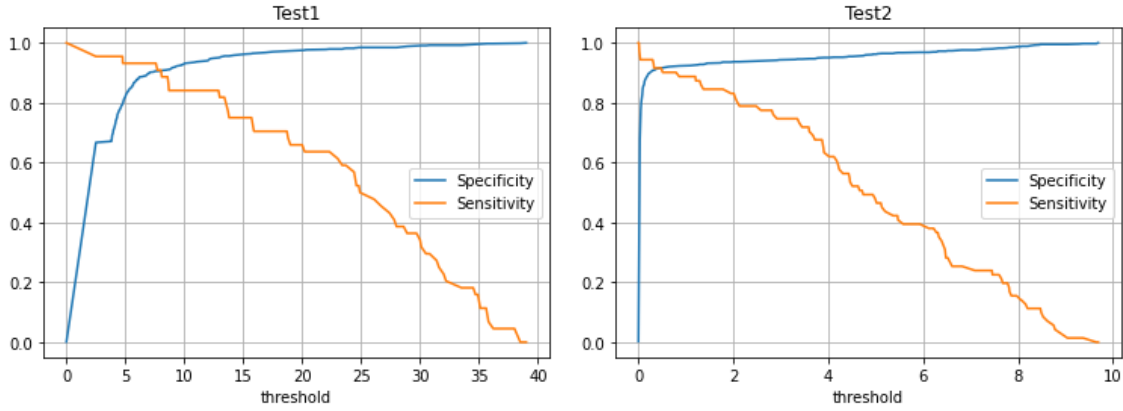


Figure 2: Sensitivity and specificity versus threshold for Test 1 (left) and Test 2 (right).

ability $P(T_p|H)$, with the only drawback that healthy people stay for a couple of days at home, maybe in an anxious state.

What cannot be accepted is instead a large false negative probability: in this case nasopharyngeal swab is not tested, and the person can spread the virus to many others. Thus, it is important to have a large sensitivity $P(T_p|D)$ (probability that the test is positive given that the person has the disease), but even more important is the probability $P(D|T_n)$ that the person has the disease given the test is negative. This last probability should be kept as small as possible.

Having assumed COVID-19 prevalence equal to 2 %, Figs. 3 and 4 respectively show versus the threshold:

1. The probability $P(D|T_p)$ that the patient is truly infected given that the test is positive and the probability $P(D|T_n)$ that the patient is infected given that the test is negative.
2. The probability $P(H|T_p)$ that the patient is healthy given that the test is positive and the probability $P(H|T_n)$ that the patient is healthy given that the test is negative.

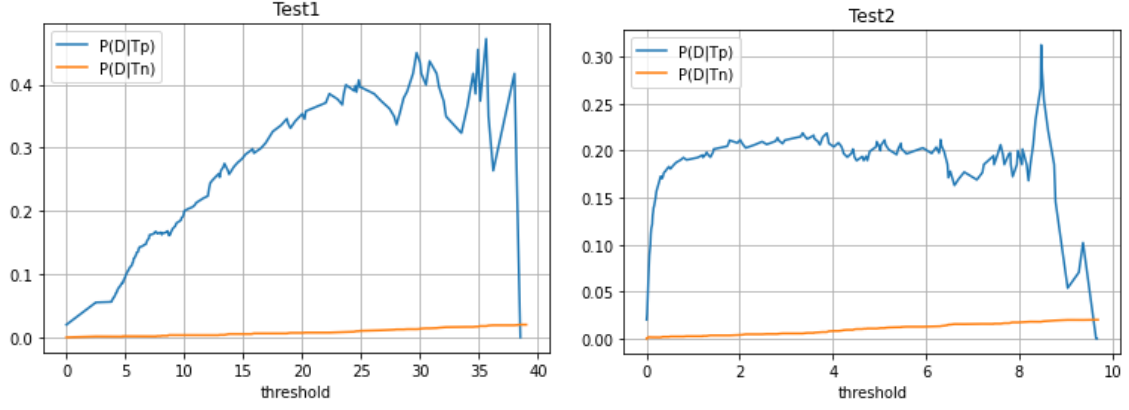


Figure 3: $P(D|T_p)$ and $P(D|T_n)$ versus threshold for Test 1 (left) and Test 2 (right).

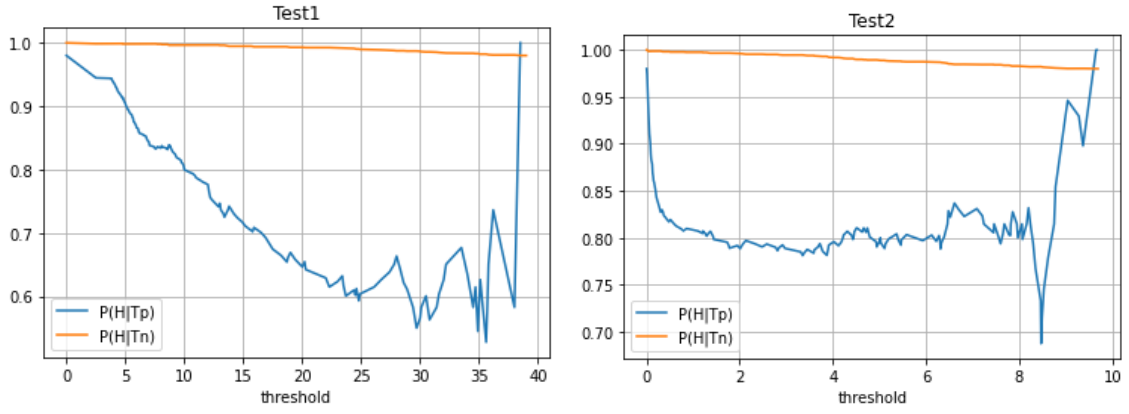


Figure 4: $P(H|T_p)$ and $P(H|T_n)$ versus threshold for Test 1 (left) and Test 2 (right).

3 Choice of the threshold

3.1 Test 1

- Sensitivity and specificity are both equal to 0.9 when the threshold is equal to 8, in which case $P(D|T_n) = 0.002$ and $P(D|T_p) = 0.162$.
- Being this sensitivity not large enough, it is convenient to decrease the threshold; we suggest a threshold equal to 4.76 for which:
 - Sensitivity $P(T_p|D) = 0.95$, false negative probability $P(T_n|D) = 0.05$.
 - Specificity $P(T_n|H) = 0.8$, false positive probability $P(T_p|H) = 0.2$.
 - $P(D|T_n) = 0.001$, $P(H|T_n) = 0.998$.
 - $P(D|T_p) = 0.086$, $P(H|T_p) = 0.913$.

3.2 Test 2

- Sensitivity and specificity are both equal to 0.9 when the threshold is equal to 0.4, in which case $P(D|T_n) = 0.001$ and $P(D|T_p) = 0.066$.
- Since this sensitivity cannot be considered sufficient, the lower threshold 0.3 is recommended, for which:
 - Sensitivity $P(T_p|D) = 0.943$, false negative probability $P(T_n|D) = 0.057$.
 - Specificity $P(T_n|H) = 0.91$, false positive probability $P(T_p|H) = 0.09$.
 - $P(D|T_n) = 0.001$, $P(H|T_n) = 0.998$.
 - $P(D|T_p) = 0.172$, $P(H|T_p) = 0.827$.

4 Conclusions

Test 1 contained outliers that were removed using the DBSCAN algorithm. After adjusting the thresholds, both tests showed approximately the same value. Test 2 shows slightly better specificity than test 1. Considering data cleaning in test 1 and the slightly better specificity in test 2, it is preferable to use test 2.

References

- [1] Ester M., Kriegel H-P., Sander J., Xu X. *A density-based algorithm for discovering clusters in large spatial databases with noise*. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining 1996 (KDD-96), pp. 226-231. AAAI Press.